

MAJORIZATION AND A SCHUR-HORN THEOREM FOR POSITIVE COMPACT OPERATORS, THE NONZERO KERNEL CASE

JIREH LOREAUX* AND GARY WEISS**

ABSTRACT. Schur-Horn theorems focus on determining the diagonal sequences obtainable for an operator under all possible basis changes, formally described as the range of the canonical conditional expectation of its unitary orbit.

Following a brief background survey, we prove an infinite dimensional Schur-Horn theorem for positive compact operators with infinite dimensional kernel, one of the two open cases posed in [KW10]. There, Kaftal and Weiss characterized the diagonals of operators in the unitary orbits for finite rank or zero kernel positive compact operators. Here we show how the characterization problem depends on the dimension of the kernel when it is finite- or infinite-dimensional.

We obtain exact majorization characterizations of the range of the canonical conditional expectation of the unitary orbits of positive compact operators with infinite-dimensional kernel, unlike the approximate characterizations of Arveson and Kadison in [AK06], but extending the exact characterizations of Gohberg and Markus [GM64] and [KW10].

Recent advances in this subject and related subjects like traces on ideals show the relevance of new kinds of sequence majorization as in the work of Kaftal-Weiss (e.g., strong majorization and another majorization similar to what here we call p -majorization), and of Kalton-Sukochev (e.g., uniform Hardy-Littlewood majorization in [KS08]), and of Bownik-Jasper (e.g., Riemann and Lebesgue majorization in [BJ13],[BJ12]). Likewise key tools here are new kinds of majorization, which we call p - and approximate p -majorization ($0 \leq p \leq \infty$).

1. INTRODUCTION

The Schur-Horn Theorem in finite matrix theory characterizes the diagonals of a self-adjoint $n \times n$ matrix in terms of its eigenvalues. In particular, if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of a self-adjoint matrix counting multiplicity, then its diagonal sequence d_1, \dots, d_n has the following relationship with its eigenvalues:

$$\sum_{i=1}^m d_i^* \leq \sum_{i=1}^m \lambda_i^*, \quad \text{for } 1 \leq m \leq n, \quad \text{and} \quad \sum_{i=1}^n d_i = \sum_{i=1}^n \lambda_i,$$

where d^*, λ^* are any monotone decreasing rearrangements of d, λ . This relationship between the sequences $d, \lambda \in \mathbb{R}^n$ is called majorization and is historically denoted

2010 *Mathematics Subject Classification.* Primary 47B07, 47B65, 15B51; Secondary 47A10, 47A12, 47L07.

Key words and phrases. Schur-Horn Theorem, majorization, diagonals, stochastic matrices.

*Partially supported by funds from the Charles Phelps Taft Research Center.

**Partially supported by Simons Foundation Collaboration Grant for Mathematicians #245014 and the Charles Phelps Taft Research Center.

by $d \prec \lambda$. Schur proved this diagonal-eigenvalue relationship in [Sch23] and Horn [Hor54] proved the converse. That is, Horn proved that given $d \prec \lambda$, there exists a self-adjoint $n \times n$ matrix with eigenvalue sequence λ and diagonal sequence d .

To modernize this perspective, let H denote a Hilbert space (finite or separable infinite dimensional) and fix an orthonormal basis $\{e_n\}_{n=1}^N$ for H ($1 \leq N \leq \infty$). Denote by \mathcal{D} the abelian algebra of diagonal operators (the canonical atomic masa of $B(H)$) corresponding to the basis $\{e_n\}_{n=1}^N$ and \mathcal{D}_{sa} the self-adjoint operators in \mathcal{D} . Given an operator $X \in B(H)$, we denote by $E(X)$ the diagonal operator having as its diagonal the main diagonal of X (i.e., $E : B(H) \rightarrow \mathcal{D}$ is the canonical faithful normal trace-preserving conditional expectation). Let $\mathcal{U}(H)$ be the full unitary group of $B(H)$, and given an operator $X \in B(H)$, $\mathcal{U}(X)$ will denote the orbit of X under $\mathcal{U}(H)$ acting by conjugation $X \mapsto UXU^*$. With this notation we can state the classical Schur-Horn Theorem ([Sch23], [Hor54]) in a form which translates naturally to the infinite dimensional case (see for instance Theorems 1.8, 1.9 and Corollary 3.5).

Theorem 1.1 (Classical Schur-Horn Theorem [Sch23, Hor54]). *Let H be a finite dimensional complex Hilbert space and \mathcal{D} a masa of $B(H)$ ($\cong M_n(\mathbb{C})$) relative to a fixed basis corresponding to \mathcal{D}) with conditional expectation $E : B(H) \rightarrow \mathcal{D}$. Then for any self-adjoint operator $A \in B(H)$,*

$$E(\mathcal{U}(A)) = \{B \in \mathcal{D}_{sa} \mid \lambda(B) \prec \lambda(A)\},$$

where $\lambda(A), \lambda(B)$ denote the eigenvalue sequences of A, B counting multiplicity.

Since the advent of the Schur-Horn theorem, there has been significant progress towards developing infinite dimensional analogues. This was perhaps started by the work of Markus [Mar64] and Gohberg and Markus [GM64], but more recently the topic was revived by A. Neumann in [Neu99]. However, Neumann studied an approximate Schur-Horn phenomenon, i.e., the operator norm closure of the diagonals of bounded self-adjoint operators (equivalently, the ℓ^∞ -norm closure of the diagonal sequences), which Arveson and Kadison deemed too coarse a closure [AK06, Introduction ¶3]. Instead, they studied the expectation of the trace norm closure of the unitary orbit of a trace class operator and then proved Schur-Horn analogues for trace class operators in $B(H)$ (type I_∞ factors). They also formulated a Schur-Horn conjecture for type II_1 factors, but discussion of this topic is outside the scope of this paper. For work on II_1 and II_∞ factors, see the work of Argerami and Massey [AM07, AM08, AM13], and a recent unpublished work Ravichandran [Rav12].

Basic notation for this paper. For a set S , let $|S|$ denote its cardinality. Let c_0^+ denote the cone of nonnegative sequences converging to zero and c_0^* the cone of nonnegative decreasing sequences converging to zero. For a sequence $\xi \in c_0^+$, let $\xi^* \in c_0^*$ denote the monotonicization of ξ , or rather the monotonicization of $\xi|_{\text{supp } \xi}$ when ξ is not finitely supported. That is, ξ_j^* denotes the j -th largest element of ξ . Notice that if ξ is finitely supported, then ξ^* ends in zeros. However, if ξ has infinite support, then ξ^* has *no* zeros, and in this case the monotonicization ξ^* does not reflect the zeros of ξ nor their multiplicity.

The following Definition 1.2 agrees with most of the literature, but it is a departure from that of [KW10] which did not include this equality condition. When they needed an equality-like condition, they used instead the more restrictive Definition 1.3 of strong majorization.

Definition 1.2. Let $\xi, \eta \in c_0^+$. One says that ξ is *majorized* by η , denoted $\xi \prec \eta$, if for all $n \in \mathbb{N}$,

$$\sum_{j=1}^n \xi_j^* \leq \sum_{j=1}^n \eta_j^* \quad \text{and} \quad \sum_{j=1}^{\infty} \xi_j = \sum_{j=1}^{\infty} \eta_j.$$

Definition 1.3 ([KW10, Definition 1.2]). Let $\xi, \eta \in c_0^+$. One says that ξ is *strongly majorized* by η , denoted $\xi \prec_s \eta$, if for all $n \in \mathbb{N}$,

$$\sum_{j=1}^n \xi_j^* \leq \sum_{j=1}^n \eta_j^* \quad \text{and} \quad \liminf_n \left\{ \sum_{j=1}^n (\eta_j^* - \xi_j^*) \right\} = 0.$$

Note that when $\xi \prec \eta \in \ell^1$, so is $\xi \in \ell^1$, and in this η -summable case, majorization *as defined above* in Definition 1.2 is equivalent to strong majorization in Definition 1.3. However, in the nonsummable case, the latter is clearly a stronger constraint than the former. Strong majorization is not an essential tool in the main theorems of this paper, but we thought it important to emphasize the distinction between our definition of majorization and that of Kaftal-Weiss just stated above.

The reason for our Definition 1.2 departure from that of Kaftal-Weiss is for convenience, efficiency of notation and unification of cases. This notation allows us to state in a more unified way the results for both trace-class *and* non trace-class operators simultaneously without splitting the conclusions into cases (compare Theorem 1.7 to the two cases in [KW10, Corollary 5.4]).

Recent History. In [KW10], Kaftal and Weiss provided an exact extension of the Schur-Horn Theorem to positive compact operators, i.e., precise results without taking closures of any kind. That is, in terms of majorization they characterize precisely the expectation of the unitary orbit of strictly positive compact operators and the expectation of the partial isometry orbit for all positive compact operators. And they ask for but leave as an open question a characterization of the expectation of the unitary orbit of positive compact operators with nonzero kernel.

To describe the subject requires some traditional preliminaries. The range projection R_A for operators $A \in B(H)$ is the orthogonal projection onto $\overline{\text{ran } A} = \ker^\perp A^*$. Thus for a self-adjoint operator A , R_A^\perp is the projection onto $\ker A$ and hence $\text{Tr } R_A^\perp = \dim \ker A$, and $\text{Tr } R_A = \text{rank } A$. Throughout this paper we opt for using $\text{Tr } R_A$ and $\text{Tr } R_A^\perp$ instead of $\dim \ker A$ and $\text{rank } A$.

Definition 1.4. Given an operator $A \in B(H)$, the *partial isometry orbit* of A is the set

$$\mathcal{V}(A) = \{VAV^* \mid V \in B(H), V^*V = R_A \vee R_{A^*}\}.$$

Notice this extends to partial isometries the standard notation of unitary orbits $\mathcal{U}(A) = \{UAU^* \mid \text{unitary } U \in \mathcal{U}(H)\}$.

Stochastic matrices play a central role in this subject due to the following definition and lemma.

Definition 1.5. A matrix P with positive entries is called

- *substochastic* if its row and column sums are bounded by 1;
- *column-stochastic* if it is substochastic and its column sums equal 1;
- *row-stochastic* if it is substochastic and its row sums equal 1;
- *doubly stochastic* if it is row- and column-stochastic;

- *unistochastic* if it is the Schur-product of a unitary matrix with its complex conjugate
(the *Schur-product* of two matrices $A = (a_{ij})$ and $B = (b_{ij})$ is the matrix $(a_{ij}b_{ij})$, that is, it is the entrywise product of A, B);
- *orthostochastic* if it is the Schur-square of an orthogonal matrix, i.e., unitary with real entries.

And the connection between expectations of orbits and stochastic matrices is

Lemma 1.6 ([KW10, Lemmas 2.3, 2.4]). *Let $\xi, \eta \in \ell^\infty$ and for any contraction $L = (L_{ij}) \in B(H)$, let $Q_{ij} := |L_{ij}|^2$ for all i, j . Then*

$$\xi = Q\eta \quad \text{if and only if} \quad \text{diag } \xi = E(L \text{diag } \eta L^*).$$

Furthermore,

- (i) Q is substochastic;
- (ii) L is an isometry if and only if Q is column-stochastic;
- (iii) L is a co-isometry (isometry adjoint) if and only if Q is row-stochastic;
- (iv) L is unitary if and only if Q is unistochastic;
- (v) L is orthogonal if and only if Q is orthostochastic.

For completeness we repeat the straightforward short proof.

Proof. Given $\xi, \eta \in \ell^\infty$, notice that for any $n \in \mathbb{N}$,

$$\begin{aligned} \langle E(L \text{diag } \eta L^*)e_n, e_n \rangle &= \langle L \text{diag } \eta L^*e_n, e_n \rangle \\ &= \left\langle (\text{diag } \eta) \sum_{j=1}^{\infty} \bar{L}_{nj}e_j, \sum_{k=1}^{\infty} \bar{L}_{nk}e_k \right\rangle \\ &= \left\langle \sum_{j=1}^{\infty} \eta_j \bar{L}_{nj}e_j, \sum_{k=1}^{\infty} \bar{L}_{nk}e_k \right\rangle \\ &= \sum_{j=1}^{\infty} |L_{nj}|^2 \eta_j = (Q\eta)_n. \end{aligned}$$

Notice now that

$$(1.1) \quad \sum_{j=1}^{\infty} Q_{ij} = \sum_{j=1}^{\infty} L_{ij}L_{ji}^* = \langle LL^*e_i, e_i \rangle = \|L^*e_i\|^2 \leq 1 \quad \text{for every } i,$$

and similarly

$$(1.2) \quad \sum_{i=1}^{\infty} Q_{ij} = \|Le_j\|^2 \leq 1 \quad \text{for every } j.$$

- (i) Immediate from (1.1) and (1.2).
- (ii) If L is an isometry, then it is immediate from (i) and the equality cases of (1.2) that Q is column-stochastic. Conversely assume that Q is column-stochastic and hence $\|Le_j\| = 1$ for all j by (1.2). Then $\langle L^*Le_j, e_j \rangle = 1$ for all j and thus it follows that $E(I - L^*L) = 0$. Since E is faithful and $I - L^*L \geq 0$ because L is a contraction by hypothesis, it follows that $L^*L = I$.
- (iii) Apply (ii) to L^* .
- (iv) Immediate from (ii) and (iii).
- (v) Immediate from (iv) and the fact that L has real entries. ■

Many of the results in [KW10] are stated and proved in terms of these stochastic matrices. We state here some of their theorems more relevant to this study.

Theorem 1.7 ([KW10, Corollary 5.4]). *If $\xi, \eta \in c_o^*$, then*

$$\xi = Q\eta \text{ for some orthostochastic matrix } Q \iff \xi \prec \eta.$$

As an integration of the summable and nonsummable cases, above Theorem 1.7 as stated is an example of the convenience afforded by our definition of majorization in contrast with that of [KW10, Corollary 5.4], where the summable and non-summable cases are combined here under the new notation.

Using these tools, Kaftal and Weiss go on to prove an infinite dimensional analog of the Schur-Horn Theorem for partial isometry orbits. This includes the unitary orbits for strictly positive compact operators (see the next two theorems).

Theorem 1.8 ([KW10, Proposition 6.4]). *Let $A \in K(H)^+$. Then*

$$E(\mathcal{V}(A)) = \{B \in \mathcal{D} \cap K(H)^+ \mid s(B) \prec s(A)\}.$$

Again, comparing this statement of the theorem with [KW10, Proposition 6.4], one sees the convenience of defining majorization as in Definition 1.2 as opposed to [KW10, Definition 1.2].

Focusing on the partial isometry orbit as opposed to the unitary orbit in the above theorem sidesteps the effects of the dimension of the kernel of the operator A . In this way, this theorem avoids the difficulties that lie therein. A similar situation appeared in [AK06] when they studied $E(\overline{\mathcal{U}(A)}^{\|\cdot\|_1})$ for positive trace class operators A , which does not involve kernel dimension considerations. In addition, they showed that $\mathcal{V}(A) = \overline{\mathcal{U}(A)}^{\|\cdot\|_1}$. Since neither of these objects encodes the dimension of the kernel of A and because they coincide for trace class operators, $\mathcal{V}(A)$ is a natural substitute for $\overline{\mathcal{U}(A)}^{\|\cdot\|_1}$ for positive compact operators A (outside the trace class) in the context of Schur-Horn theorems.

However, the question of precisely what is $E(\mathcal{U}(A))$ for all $A \in K(H)^+$ was only partially answered in [KW10]. In particular, it was answered when A has finite rank or when $R_A = I$ (that is, when A is strictly positive). When A has finite rank, $\mathcal{U}(A) = \mathcal{V}(A)$ (see Theorem 2.4, proof case 1), and so is covered by Theorem 1.8. For the case when $R_A = I$, they have

Theorem 1.9 ([KW10, Proposition 6.6]). *Let $A \in K(H)^+$ with $R_A = I$. Then*

$$E(\mathcal{U}(A)) = E(\mathcal{V}(A)) \cap \{B \in \mathcal{D} \mid R_B = I\}.$$

Our main contribution. Theorem 1.9 left open the case when A has infinite rank and nonzero kernel. We attempt here to close this gap. In particular, we characterize $E(\mathcal{U}(A))$ when A has infinite dimensional kernel. When A has finite dimensional kernel, we give a necessary condition for membership in $E(\mathcal{U}(A))$ (which we conjecture is also sufficient), and we give a sufficient condition for membership in $E(\mathcal{U}(A))$ (which we know not to be necessary when $0 < \text{Tr } R_A^{-1} < \infty$, see Example 2.6 below which also appears in [KW10, Proposition 6.10, Example 6.11]). These main results are embodied in Theorems 2.4, 3.4 and Corollary 3.5. Both of these membership conditions involve new kinds of majorization, which here we call p -majorization and herein our brand new approximate p -majorization (for $0 \leq p \leq \infty$, Definitions 2.2 and 3.1 below). There is a natural hierarchy of these new types of majorization which the following diagram describes succinctly. All of these implications are

natural (see the discussions following Definitions 2.2 and 3.1) except the two corresponding to the dashed arrows, which are handled in Proposition 2.7 and are only applicable when both sequences in question are in $c_0^+ \setminus \ell^1$. A linear interpretation of the diagram in Figure 1 is presented below in Figure 2.

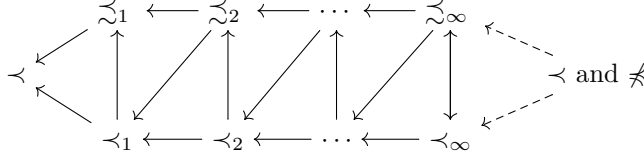


FIGURE 1. Hierarchy of majorization

$$\prec \leftarrow \sim_1 \leftarrow \prec_1 \leftarrow \sim_2 \leftarrow \cdots \leftarrow \sim_\infty \longleftrightarrow \prec_\infty \dashleftarrow (\prec \text{ and } \not\prec)$$

FIGURE 2. Linear hierarchy of majorization

2. p -MAJORIZATION (SUFFICIENCY)

A result which was known to Kaftal and Weiss, and almost certainly to others, is a necessary condition for membership in the expectation of the unitary orbit of a positive operator (see [KW10, Proof of Lemma 6.9]). Namely, that the dimensions of the kernels of operators in the range of the expectation of the unitary orbit of a positive operator cannot increase from the dimension of the kernel of the operator itself.

Proposition 2.1. *If $A \in B(H)^+$ and $B \in E(\mathcal{U}(A))$, then $\ker B \subseteq \ker(UAU^*)$ for some $U \in \mathcal{U}(H)$, and hence also $\text{Tr}(R_B^\perp) \leq \text{Tr}(R_A^\perp)$.*

Vector proof. Since $B \in E(\mathcal{U}(A))$, $B = E(UAU^*)$ for some unitary $U \in \mathcal{U}(H)$, and since $\text{Tr}(R_A^\perp) = \dim \ker A = \dim \ker UAU^* = \text{Tr}(R_{UAU^*}^\perp)$, the required trace inequality follows from the inclusion $\ker B \subseteq \ker UAU^*$, which itself follows from $A \geq 0$ and

$$\begin{aligned} \ker B &= \ker E(UAU^*) \\ &= \overline{\text{span}}\{e_n \mid \langle E(UAU^*)e_n, e_n \rangle = 0\} \\ &= \overline{\text{span}}\{e_n \mid \langle UAU^*e_n, e_n \rangle = 0\} \\ &= \overline{\text{span}}\{e_n \mid \|(UAU^*)^{1/2}e_n\|^2 = 0\} \\ &\subseteq \ker(UAU^*)^{1/2} = \ker UAU^* \end{aligned} \quad \blacksquare$$

Projection proof. Since $A \geq 0$, R_A^\perp is the largest projection P so that $PAP = 0$. Since $B \in E(\mathcal{U}(A))$, there is some unitary $U \in \mathcal{U}(H)$ so that $B = E(UAU^*)$.

Notice also that $R_B^\perp \in \mathcal{D}$ since $B \in \mathcal{D}$. Therefore,

$$\begin{aligned} E(U(U^* R_B^\perp U)A(U^* R_B^\perp U)U^*) &= E(R_B^\perp U A U^* R_B^\perp) \\ &= R_B^\perp E(U A U^*) R_B^\perp = R_B^\perp B R_B^\perp = 0. \end{aligned}$$

Then E is faithful and $U(U^* R_B^\perp U)A(U^* R_B^\perp U)U^*$ is positive, it is zero, and thus conjugating by U^* shows $(U^* R_B^\perp U)A(U^* R_B^\perp U) = 0$. By the maximality of R_A^\perp , $U^* R_B^\perp U \leq R_A^\perp$ and therefore

$$\mathrm{Tr}(R_B^\perp) = \mathrm{Tr}(U^* R_B^\perp U) \leq \mathrm{Tr}(R_A^\perp). \quad \blacksquare$$

Before we proceed with our analysis, we need to introduce next a concept similar to [KW10, Definition 6.8(ii)] which here we call p -majorization. Roughly speaking, it is majorization along with eventual p -expanded majorization. And this led us to the definition below of ∞ -majorization which is both new and fruitful.

Definition 2.2. Given $\xi, \eta \in c_0^+$ and $0 \leq p < \infty$, we say that ξ is p -majorized by η , denoted $\xi \prec_p \eta$, if $\xi \prec \eta$ and there exists an $N_p \in \mathbb{N}$ such that for all $n \geq N_p$, we have the inequality

$$\sum_{k=1}^{n+p} \xi_k \leq \sum_{k=1}^n \eta_k.$$

And ∞ -majorization, denoted $\xi \prec_\infty \eta$, means $\xi \prec_p \eta$ for all $p \in \mathbb{N}$.

Note that $\xi \prec_0 \eta$ is precisely the statement that $\xi \prec \eta$ (recall Definition 1.2, which includes equality of the sums). One also observes that if $\xi \prec_p \eta$ and $p' \leq p$, then $\xi \prec_{p'} \eta$ (we use often the special case that p -majorization implies 0-majorization, i.e., majorization). For this reason, $\xi \prec_p \eta$ for infinitely many p is equivalent to $\xi \prec_p \eta$ for all $p < \infty$, in which case we say that ξ is ∞ -majorized by η and we write $\xi \prec_\infty \eta$.

One should also take note that p -majorization is actually strictly stronger than p' -majorization when $p' < p$. That is, there exist sequences $\xi, \eta \in c_0^+$ for which $\xi \prec_{p'} \eta$ but $\xi \not\prec_p \eta$. From the remarks of the previous paragraph, it suffices to exhibit ξ, η when $p = p' + 1$. To produce such sequences, start with any $0 < \eta \in c_0^+$ and define

$$\xi^{(p)} := \underbrace{\langle \eta_1/p, \dots, \eta_1/p \rangle}_{p \text{ times}}, \eta_2, \eta_3, \dots.$$

Even though $\xi^{(p)}$ is not necessarily monotone, it is not difficult to verify that $\xi^{(p)} \prec_{p'} \eta$ but $\xi^{(p)} \not\prec_p \eta$.

Remark 2.3. Because $\mathcal{V}(A) \supseteq \mathcal{U}(A)$, Theorem 1.8 for $A \in K(H)^+$ implies that $s(B) \prec s(A)$ is a necessary condition for $B \in E(\mathcal{U}(A))$; Proposition 2.1 shows $\mathrm{Tr} R_B^\perp \leq \mathrm{Tr} R_A^\perp$ is another necessary condition for membership in $E(\mathcal{U}(A))$ and with majorization is equivalent to membership in $E(\mathcal{U}(A))$ when $\mathrm{Tr} R_A^\perp = 0$ by Theorem 1.9. It was natural in [KW10] to ask how the role of majorization is impacted by the dimension of these kernels. In particular here we enhance this program by asking, what role, if any, $p = \mathrm{Tr} R_A^\perp - \mathrm{Tr} R_B^\perp$ plays in relating membership in $E(\mathcal{U}(A))$ to majorization. And when this difference is undefined, $p = 0$ is the minimal p for which $\mathrm{Tr} R_A^\perp \leq \mathrm{Tr} R_B^\perp + p$. This is the strategy that guided our program.

The result below appears in [KW10] as Lemma 6.9 for $p < \infty$, the proof of which utilizes orthostochastic matrices. We provide a different proof which instead utilizes expectations of unitary orbits because it leads to a straightforward extension to the $p = \infty$ case. See Remark 2.5 for when an orthostochastic matrix can be produced to implement the construction. But we do not have a complete characterization for this orthostochasticity case.

Theorem 2.4. *Let $A, B \in K(H)^+$, $B \in \mathcal{D}$ and $\text{Tr}(R_B^\perp) \leq \text{Tr}(R_A^\perp)$. If for some $0 \leq p \leq \infty$, $s(B) \prec_p s(A)$ and $\text{Tr}(R_A^\perp) \leq \text{Tr}(R_B^\perp) + p$, then $B \in E(\mathcal{U}(A))$.*

Proof. Case 1: A has finite rank.

In this case $\mathcal{U}(A) = \mathcal{V}(A)$, even if A is not necessarily self-adjoint. Indeed, the elements of $\mathcal{V}(A)$, by Definition 1.4, have the form VAV^* for some partial isometry for which $V^*V = R_A \vee R_{A^*} (= P)$, so $PA = AP = A$. Then $VAV^* = (VP)A(VP)^*$ and VP is also a partial isometry with $(VP)^*(VP) = P$. But this partial isometry VP is finite rank since R_A and R_{A^*} and hence also P are finite rank, and so VP can be extended to a unitary U for which $VAV^* = (VP)A(VP)^* = UAU^*$. This shows $\mathcal{V}(A) \subseteq \mathcal{U}(A)$, and hence equality. That the conclusion of Theorem 2.4 holds in this case is then covered by Theorem 1.8, since $A \geq 0$, $\mathcal{U}(A) = \mathcal{V}(A)$, and since $s(B) \prec_p s(A)$ implies $s(B) \prec s(A)$.

In the proof of Case 1, we only used the facts that A had finite rank and $s(B) \prec s(A)$. Although not needed in this case, the other hypotheses hold automatically and for edification we explain why. Even though $s(B) \prec_p s(A)$ for some $p \geq 0$ implies $s(B) \prec s(A)$, one has the stronger converse: when $s(A)$ has finite support, $s(B) \prec s(A)$ implies $s(B) \prec_p s(A)$ for every $p \geq 0$, i.e., $s(B) \prec_\infty s(A)$. Indeed, let N_p be the largest index for which $s(A)$ has a nonzero value. Then for all $k \geq N_p$, we have

$$\sum_{j=1}^{k+p} s_j(B) \leq \sum_{j=1}^{k+p} s_j(A) = \sum_{j=1}^k s_j(A)$$

and therefore $s(B) \prec_p s(A)$. Since p is arbitrary, $s(B) \prec_\infty s(A)$. This shows that the second inequality in the hypotheses is satisfied for $p = \infty$ since its right-hand side is infinite. The first inequality is satisfied since $\text{Tr}(R_A^\perp) = \infty$ because A is finite rank.

Case 2: A has infinite rank and $\text{Tr}(R_B^\perp) = \text{Tr}(R_A^\perp)$.

Then $\dim \ker B = \dim \ker A$. That A has infinite rank along with both the inequality and equality parts of majorization implies that B has infinite rank. Indeed, consider the contrapositive. That is, since $s(B) \prec_p s(A)$ implies $s(B) \prec s(A)$, the latter majorization itself implies that $s(B) > 0$, otherwise equality of their infinite sums fails when A has infinite rank, in particular because $s(A) > 0$. But then $s(B)$ is an infinite nonzero sequence, so B has infinite rank.

Without loss of generality we may next assume that $A \in \mathcal{D}$ (because $\mathcal{U}(A) = \mathcal{U}(UAU^*)$ for every $U \in \mathcal{U}(H)$) and then via equality of kernel dimensions and using permutation unitaries one ensures that we may also assume

$$\mathbb{N}_{(1)} := \{n \in \mathbb{N} \mid Be_n = 0\} = \{n \in \mathbb{N} \mid Ae_n = 0\}.$$

Let $\mathbb{N}_{(2)} := \mathbb{N} \setminus \mathbb{N}_{(1)}$. Let $H_i = \ell^2(\mathbb{N}_{(i)})$ for $i = 1, 2$. Then $H = H_1 \oplus H_2$. Furthermore, if A_i, B_i denote the compressions of A, B to H_i , then one obtains: $A = A_1 \oplus A_2$, $B = B_1 \oplus B_2$, $A_1 = B_1 = 0$, $s(A_2) = s(A) \prec s(B) = s(B_2)$,

and $R_{A_2} = R_{B_2} = I_{H_2}$. So by Theorem 1.9, there is a unitary U_2 on H_2 so that $B_2 = E(U_2 A_2 U_2^*)$. Letting $U = I_{H_1} \oplus U_2$, one has that $B = E(U A U^*)$.

Case 3: A has infinite rank and $N = \text{Tr}(R_B^\perp) < \text{Tr}(R_A^\perp)$. (The most complicated case)

Since $\text{Tr } R_A^\perp \leq \text{Tr } R_B^\perp + p$, if necessary, by passing to a smaller p , we may assume $\text{Tr}(R_A^\perp) = \text{Tr}(R_B^\perp) + p$, even if $\text{Tr } R_A^\perp = \infty$ in which case $p = \infty$.

Using a similar splitting argument as in Case 2, it suffices to prove Case 3 when $N = 0$ and A, B are simultaneously diagonalized. Indeed, because $\dim \ker B = \text{Tr } R_B^\perp < \text{Tr } R_A^\perp = \dim \ker A$, without loss of generality we can assume $\ker B \subseteq \ker A$. Then $H = S \oplus S^\perp$, $A = A_1 \oplus A_2$, $B = B_1 \oplus B_2$, $A_1 = 0 = B_1$, and B_2 has zero kernel, which is the $N = 0$ case for proving Theorem 2.4. That is, it suffices to assume $N = 0$ for Case 3. Since $\text{Tr } R_A^\perp = \text{Tr } R_B^\perp + p$, in this case of $N = 0$, one has $\text{Tr } R_A^\perp = p$ and since $0 = \text{Tr } R_B^\perp < \text{Tr } R_A^\perp$, one has $1 \leq p \leq \infty$.

To prove the case $N = 0$ we employ another splitting. First let $\{e_j\}_{j=1}^\infty$ denote the basis that diagonalizes A, B , and then assign them different names, that is, let $\{f_j\}_{j=1}^p$ be the collection of e_j 's such that $(Ae_j, e_j) = 0$. Since A is diagonalized with respect to the basis $\{e_j\}_{j=1}^\infty$ the collection $\{f_j\}_{j=1}^p$ forms an orthonormal basis for $\ker A$. Let $\{g_j\}_{j=1}^\infty$ consist of the remainder of the set $\{e_j\}_{j=1}^\infty$, which means $\{g_j\}_{j=1}^\infty$ is a basis for $\ker^\perp A$. Then $H = H_1 \oplus H_2$, where $H_1 = \ker A$ and $H_2 = \ker^\perp A$. Let $\mathbb{N}_p := \{n \in \mathbb{N} \mid n \leq p\}$ and then define $\text{diag}_{H_1 \oplus H_2} : \ell^\infty(\mathbb{N}_p) \times \ell^\infty(\mathbb{N}) \rightarrow \mathcal{D}$ by

$$(2.1) \quad \text{diag}_{H_1 \oplus H_2} \langle \varphi, \rho \rangle = \left(\begin{array}{ccc|ccc} \varphi_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \varphi_p & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & \rho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \rho_2 & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \ddots \end{array} \right)$$

Then because of the way in which we chose $\{f_j\}_{j=1}^p$ and $\{g_j\}_{j=1}^\infty$ and because $N = 0$, if we let $\eta' := s(A)$ and $\xi = s(B)$ then A, B are

$$A = \text{diag}_{H_1 \oplus H_2} \langle \mathbf{0}, \eta' \rangle \quad \text{and} \quad B = \text{diag } \xi,$$

where $\mathbf{0} \in \ell^\infty(\mathbb{N}_p)$ is the zero sequence and by hypothesis $\xi \prec_p \eta'$.

The heuristic idea of the proof is the following in descriptive informal language. First construct a sequence ξ' which is a sparsely compressed version of ξ but sufficient to retain majorization by η , i.e., $\xi' \prec \eta$. Next apply Case 2 to obtain a special unitary U for which $E(U(\text{diag } \langle \mathbf{0}, \eta \rangle)U^*) = \text{diag } \langle \mathbf{0}, \xi' \rangle$. Finally, apply another unitary to decompress the diagonal $\langle \mathbf{0}, \xi' \rangle$ to the diagonal ξ .

Now inductively choose sequences of nonnegative integers $\{N_m\}_{m=1}^p$ and $\{N'_m\}_{m=0}^p$ with the following properties:

- (i) $0 = N'_0 < N'_0 + 1 < N_1 < N'_1 < N'_1 + 1 < N_2 < N'_2 < \cdots$ if $p = \infty$,
(or $\cdots < N_p < N'_p$ if $p < \infty$).
- (ii) For each $m \in \mathbb{N}_p$, whenever $n \geq N_m - (m - 1)$, one has

$$\sum_{k=1}^{n+m} \xi_k \leq \sum_{k=1}^n \eta'_k.$$

- (iii) $\xi_{N_m} + \xi_{N'_m} \leq \xi_{N_m-1}$ for each $m \in \mathbb{N}_p$.

For transparency and brevity, we only loosely describe the construction of the pair of sequences $\{N_m\}_{m=1}^p$ and $\{N'_m\}_{m=0}^p$. The construction proceeds in pairs, N_m, N'_m . We may choose N_m to satisfy property (ii) since $\xi \prec_p \eta'$ and hence $\xi \prec_m \eta'$ because $m \leq p$. For this we use the fact that property (ii) is an eventual property in the sense that if it holds for some N_m it holds for any larger N_m . Moreover, because $\xi > 0$ and $\xi \downarrow 0$, ξ has infinitely many strictly decreasing jumps, i.e., for infinitely many j one has $\xi_j \not\leq \xi_{j-1}$. If necessary, increase N_m so that it satisfies this condition. Then since $\xi \rightarrow 0$, we may choose N'_m to satisfy property (iii). To construct the entire pair of sequences, simply iterate this process while simultaneously ensuring $N_{m+1} > N'_m + 1$, which we can guarantee because property (ii) is an eventual property.

Next since $N'_0 < N_m < N'_m < N_{m+1} < N'_{m+1}$ for all $1 \leq m < p$, one has for $p = \infty$

$$\begin{aligned} N'_0 + 1 = 1 &\leq N_1 < N'_1 \leq N_2 - 1 < N'_2 - 1 \leq N_3 - 2 < N'_3 - 2 \leq \cdots \\ \cdots &\leq N_m - (m-1) < N'_m - (m-1) \leq N_{m+1} - m < N'_{m+1} - m \leq \cdots \\ &\left(\cdots \leq N_p - (p-1) < N'_p - (p-1) \quad \text{for } p < \infty \right). \end{aligned}$$

When $p < \infty$, if we set $N'_{p+1} = \infty$ for convenience of notation, then regardless of whether $p < \infty$ or $p = \infty$ these inequalities partition

$$\mathbb{N} = \bigsqcup_{m=0}^p [N'_m - (m-1), N'_{m+1} - m)$$

with each $N_m - (m-1) \in [N'_{m-1} - (m-2), N'_m - (m-1))$ and $m \in \mathbb{N}_p$.

Next define the sequence ξ' which shifts and alters ξ at one point in each $[N'_{m-1} - (m-2), N'_m - (m-1))$: for each $m \in \mathbb{N}_p$ (or for each $m \in \mathbb{N}_{p+1}$ if $p < \infty$, in which case the last interval is $[N'_p - (p-1), \infty)$), set

$$\xi'_k = \begin{cases} \xi_{N_m} + \xi_{N'_m} & \text{if } k = N_m - (m-1); \\ \xi_{k+m-1} & \text{if } N'_{m-1} - (m-2) \leq k < N'_m - (m-1) \\ & \text{but } k \neq N_m - (m-1). \end{cases}$$

This partition of \mathbb{N} ensures that ξ' is well-defined. Property (iii) guarantees that ξ' is monotone decreasing. And property (ii) allows us to conclude that $\xi' \prec \eta'$ which will follow from the next claim.

Claim: For all $m \in \mathbb{N}_p$,

$$(2.2) \quad \sum_{j=1}^k \xi'_j = \sum_{j=1}^{k+m-1} \xi_j \quad \text{for } N'_{m-1} - (m-2) \leq k < N_m - (m-1),$$

and

$$(2.3) \quad \sum_{j=1}^k \xi'_j = \xi_{N'_m} + \sum_{j=1}^{k+m-1} \xi_j \quad \text{for } N_m - (m-1) \leq k < N'_m - (m-1).$$

The proof of this claim is by induction on m . We start with the base case $m = 1$. If $1 = N'_0 + 1 = N'_{m-1} - (m-2) \leq k < N_m - (m-1) = N_1$, then for $1 \leq j \leq k$,

$\xi'_j = \xi_{j+m-1} = \xi_{j+0} = \xi_j$. Thus

$$\sum_{j=1}^k \xi'_j = \sum_{j=1}^k \xi_j = \sum_{j=1}^{k+m-1} \xi_j,$$

verifying equation (2.2) when $m = 1$. If $N_1 = N_m - (m-1) \leq k < N'_m - (m-1) = N'_1$, then for $1 \leq j \leq k$, $j \neq N_1 = N_m - (m-1)$, one has $\xi'_j = \xi_{j+m-1} = \xi_j$ and so

$$\sum_{j=1}^k \xi'_j = \xi'_{N_1} + \sum_{\substack{j=1 \\ j \neq N_1}}^k \xi'_j = \xi_{N_1} + \xi_{N'_1} + \sum_{\substack{j=1 \\ j \neq N_1}}^k \xi_j = \xi_{N'_1} + \sum_{j=1}^k \xi_j,$$

verifying equation (2.3) when $m = 1$.

For the inductive step, fix an arbitrary $m+1 \in \mathbb{N}_p$ and suppose equations (2.2) and (2.3) hold for m . Let $N'_m - (m-1) \leq k < N_{m+1} - m$. Then for $N'_m - (m-1) \leq j \leq k$, one has $\xi'_j = \xi_{j+(m+1)-1} = \xi_{j+m}$ and so

$$\begin{aligned} \sum_{j=1}^k \xi'_j &= \sum_{j=1}^{N'_m - (m-1) - 1} \xi'_j + \sum_{j=N'_m - (m-1)}^k \xi'_j \\ &= \xi_{N'_m} + \sum_{j=1}^{N'_m - 1} \xi_j + \sum_{j=N'_m - (m-1)}^k \xi_{j+m} \\ &\quad \text{(by (2.3) for } m\text{)} \\ &= \sum_{j=1}^{N'_m} \xi_j + \sum_{j=N'_m + 1}^{k+m} \xi_j = \sum_{j=1}^{k+(m+1)-1} \xi_j \\ &\quad \text{(by an index change),} \end{aligned}$$

which verifies equation (2.2) for $m+1$.

Next let $N_{m+1} - m \leq k < N'_{m+1} - m$. For $N_{m+1} - m < j \leq k$ one has $\xi'_j = \xi_{j+m}$, and so

$$\begin{aligned} \sum_{j=1}^k \xi'_j &= \sum_{j=1}^{(N_{m+1}-m)-1} \xi'_j + \xi'_{N_{m+1}-m} + \sum_{j=(N_{m+1}-m)+1}^k \xi'_j \\ &= \sum_{j=1}^{N_{m+1}-1} \xi_j + \left(\xi_{N_{m+1}} + \xi_{N'_{m+1}} \right) + \sum_{j=(N_{m+1}-m)+1}^k \xi_{j+m} \\ &\quad \text{(by (2.2) for } m+1\text{)} \\ &= \sum_{j=1}^{N_{m+1}} \xi_j + \xi_{N'_{m+1}} + \sum_{j=N_{m+1}+1}^{k+m} \xi_j \\ &\quad \text{(by an index change)} \\ &= \xi_{N'_{m+1}} + \sum_{j=1}^{k+(m+1)-1} \xi_j, \end{aligned}$$

which proves equation (2.3) for $m+1$. By induction we have proven the claim.

When $p < \infty$, the last interval requires separate consideration. That is, if $N'_p - (p-1) \leq k < N'_{p+1} - p = \infty$ and $N'_p - (p-1) \leq j \leq k$ one has $\xi'_j = \xi_{j+p}$ and so

$$(2.4) \quad \sum_{j=1}^k \xi'_j = \sum_{j=1}^{k+p} \xi_j \quad \text{for } k \in [N'_p - (p-1), \infty)$$

because

$$\begin{aligned} \sum_{j=1}^k \xi'_j &= \sum_{j=1}^{N'_p - (p-1) - 1} \xi'_j + \sum_{j=N'_p - (p-1)}^k \xi'_j \\ &= \sum_{j=1}^{N'_p - 1} \xi_j + \xi_{N'_p} + \sum_{j=N'_p - (p-1)}^k \xi_{j+p} \\ &\quad \text{(by (2.3) for } m = p) \\ &= \sum_{j=1}^{N'_p} \xi_j + \sum_{j=N'_p + 1}^{k+p} \xi_j = \sum_{j=1}^{k+p} \xi_j \\ &\quad \text{(by an index change).} \end{aligned}$$

It is now simple to prove that $\xi' \prec \eta'$ using condition (ii). Indeed, notice that for $m \in \mathbb{N}_p$, if $N'_{m-1} - (m-2) \leq k < N'_m - (m-1)$ and since also $m-1 \in \mathbb{N}_p \cup \{0\}$ and $\xi \prec \eta'$, and considering separately the cases $m = 1$ and $m > 1$, one has

$$(2.5) \quad \sum_{j=1}^k \xi'_j \stackrel{(2.2)}{=} \sum_{j=1}^{k+m-1} \xi_j \stackrel{\xi \prec \eta' \text{ or (ii)}}{\leq} \sum_{j=1}^k \eta'_j.$$

And for $m \in \mathbb{N}_p$, if $N'_m - (m-1) \leq k < N'_{m+1} - m$, one has

$$(2.6) \quad \sum_{j=1}^k \xi'_j \stackrel{(2.3)}{=} \xi_{N'_m} + \sum_{j=1}^{k+m-1} \xi_j \leq \sum_{j=1}^{k+m} \xi_j \stackrel{(ii)}{\leq} \sum_{j=1}^k \eta'_j.$$

Finally, if $p < \infty$ and $N'_p - (p-1) \leq k < N'_{p+1} - p = \infty$, one has

$$(2.7) \quad \sum_{j=1}^k \xi'_j \stackrel{(2.4)}{=} \sum_{j=1}^{k+p} \xi_j \stackrel{(ii)}{\leq} \sum_{j=1}^k \eta'_j.$$

Recalling that \mathbb{N} is the disjoint union of $[N'_{(m-1)} - (m-2), N'_m - (m-1))$ for $m \in \mathbb{N}_p$, (2.2–2.4) and (2.5–2.7) imply that for all $k \in \mathbb{N}$

$$\sum_{j=1}^k \xi_j \stackrel{(2.2-2.4)}{\leq} \sum_{j=1}^k \xi'_j \stackrel{(2.5-2.7)}{\leq} \sum_{j=1}^k \eta'_j.$$

Passing to the limit as $k \rightarrow \infty$ yields $\sum_{j=1}^{\infty} \xi_j = \sum_{j=1}^{\infty} \xi'_j = \sum_{j=1}^{\infty} \eta'_j$ since $\xi \prec \eta'$. Hence $\xi' \prec \eta'$.

By Case 2 applied to $A = \text{diag}_{H_1 \oplus H_2} \langle \mathbf{0}, \eta' \rangle$ and $B' = \text{diag}_{H_1 \oplus H_2} \langle \mathbf{0}, \xi' \rangle$, its proof produced a unitary U of the form $I_{H_1} \oplus W$ for which

$$E(UAU^*) = B' = \text{diag}_{H_1 \oplus H_2} \langle \mathbf{0}, \xi' \rangle.$$

Since U has this form, for all $f, f' \in \{f_j\}_{j=1}^p$ and $g \in \{g_j\}_{j=1}^\infty$ one has

$$0 = \langle UAU^* f, f' \rangle = \langle UAU^* f, g \rangle = \langle UAU^* g, f \rangle.$$

Then for all $m \in \mathbb{N}_p$ define V_m on $\text{span}\{f_m, g_{N_m-(m-1)}\}$ given by the unitary 2×2 matrix

$$(2.8) \quad V_m = \begin{pmatrix} a_m & b_m \\ -b_m & a_m \end{pmatrix}$$

where

$$a_m = \sqrt{\frac{\xi_{N'_m}}{\xi_{N_m} + \xi_{N'_m}}} \quad \text{and} \quad b_m = \sqrt{\frac{\xi_{N_m}}{\xi_{N_m} + \xi_{N'_m}}}.$$

Then $C_m = \begin{pmatrix} 0 & 0 \\ 0 & \xi'_{N_m-(m-1)} \end{pmatrix}$ is the compression of UAU^* to $\text{span}\{f_m, g_{N_m-(m-1)}\}$ and if one interprets V_m as canonically acting on this same subspace, one computes

$$(2.9) \quad V_m C_m V_m^* = \begin{pmatrix} \xi_{N_m} & * \\ * & \xi_{N'_m} \end{pmatrix}$$

because

$$\begin{aligned} V_m C_m V_m^* &= \begin{pmatrix} a_m & b_m \\ -b_m & a_m \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \xi'_{N_m-(m-1)} \end{pmatrix} \begin{pmatrix} a_m & -b_m \\ b_m & a_m \end{pmatrix} \\ &= \begin{pmatrix} 0 & b_m(\xi_{N_m} + \xi_{N'_m}) \\ 0 & a_m(\xi_{N_m} + \xi_{N'_m}) \end{pmatrix} \begin{pmatrix} a_m & -b_m \\ b_m & a_m \end{pmatrix} \\ &= \begin{pmatrix} b_m^2(\xi_{N_m} + \xi_{N'_m}) & a_m b_m(\xi_{N_m} + \xi_{N'_m}) \\ a_m b_m(\xi_{N_m} + \xi_{N'_m}) & a_m^2(\xi_{N_m} + \xi_{N'_m}) \end{pmatrix} \\ &= \begin{pmatrix} \xi_{N_m} & * \\ * & \xi_{N'_m} \end{pmatrix}. \end{aligned}$$

Next let $H' = H \ominus \text{span}\{f_m, g_{N_m-(m-1)}\}_{m=1}^p$ and let V be the unitary defined by

$$V = I_{H'} \oplus \bigoplus_{m=1}^p V_m.$$

From the above computations, one can see that the diagonal operator

$$E(VUAU^*V^*) = \Pi^*(\text{diag } \xi)\Pi,$$

for an appropriate permutation Π of the basis $\{e_n\}_{n \in \mathbb{N}}$. But conjugation by operators which permute the basis corresponding to E , in particular Π , commutes with E , and $\Pi V U$ is unitary, so $E((\Pi V U)A(\Pi V U)^*) = \text{diag } \xi$. \blacksquare

Remark 2.5 (Orthostochasticity). In the above proof, if A was already diagonalized with respect to the basis $\{e_n\}_{n \in \mathbb{N}}$, then all the unitary operators either are orthogonal with respect to this basis, or can be chosen as such. Indeed, Π and V are orthogonal, and $U = I_{H_1} \oplus U_2$ where U_2 , coming from Theorem 1.9, can be chosen to be orthogonal by [KW10, Corollary 6.1, NS(ii') and S(ii')]. A consequence of this combined with Lemma 1.6 is that if $\xi, \eta \in c_0^+$, $\xi \prec_p \eta$ and if

$$|\xi^{-1}(0)| \leq |\eta^{-1}(0)| \leq p + |\xi^{-1}(0)|,$$

then there exists an orthostochastic matrix Q for which $\xi = Q\eta$.

Example 2.6 ([KW10, Proposition 6.10, Example 6.11]). The following example, in conjunction with Lemma 1.6, will show that the condition $s(B) \prec_p s(A)$ is not necessary for $B \in E(\mathcal{U}(A))$ and $p := \text{Tr}(R_A^\perp) - \text{Tr}(R_B^\perp) < \infty$. In particular, the converse of Theorem 2.4 fails for this case, but is true when $\text{Tr} R_A^\perp - \text{Tr} R_B^\perp$ is either infinite or undefined as proved by Corollary 3.5.

A counterexample is the following. Let $\tilde{\eta} = \langle 0, \eta \rangle$ where $0 < \eta \in c_0^*$. Let Q be an orthostochastic matrix with the property that $Q_{ij} = 0$ if and only if $i > j > 1$ (so that otherwise $Q_{ij} > 0$ and rows and columns sum to one), e.g., [KW10, Example 6.11]. Then choosing $\xi := Q\tilde{\eta}$ we claim that $\xi \not\prec_1 \eta$. One can see this from the calculation

$$\begin{aligned}
 \sum_{i=1}^n \xi_i^* &\geq \sum_{i=1}^n \xi_i = \sum_{i=1}^n \sum_{j=1}^{\infty} Q_{ij} \tilde{\eta}_j \\
 &= \sum_{i=1}^n \sum_{j=2}^{\infty} Q_{ij} \eta_{j-1} \\
 &= \sum_{j=2}^n \sum_{i=1}^n Q_{ij} \eta_{j-1} + \sum_{j=n+1}^{\infty} \sum_{i=1}^n Q_{ij} \eta_{j-1} \\
 (2.10) \quad &= \sum_{j=1}^{n-1} \eta_j + \sum_{j=n+1}^{\infty} \sum_{i=1}^n Q_{ij} \eta_{j-1} > \sum_{j=1}^{n-1} \eta_j,
 \end{aligned}$$

where the latter inequality follows since $\eta > 0$.

This example can be easily extended to create similar orthostochastic examples (e.g., $\tilde{Q} = \bigoplus_1^{p-1} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \oplus Q$) for when $1 < p < \infty$ and $0 < \xi \not\prec_p \eta$, but $\xi = \tilde{Q}\tilde{\eta}$ for $\tilde{\eta} = \langle 0, \eta_1, 0, \eta_2, \dots, 0, \eta_p, \eta_{p+1}, \dots \rangle$ with p zeros. Then since \tilde{Q} is orthostochastic, by Lemma 1.6 one has $\text{diag } \xi = U(\text{diag } \tilde{\eta})U^*$ for some orthogonal matrix U with $\tilde{Q}_{ij} = |U_{ij}|^2$. Therefore p -majorization does for any $1 \leq p < \infty$ is not necessary to characterize $E(\mathcal{U}(A))$.

To verify $\xi \not\prec_p \eta$ observe that for sufficiently large n , one has

$$\begin{aligned}
 \sum_{i=1}^n \xi_i^* &= \sum_{i=1}^{p-1} \eta_i + \sum_{i=1}^{n-2p+2} (Q \langle 0, \eta_p, \eta_{p+1}, \dots \rangle^*)_i \\
 &\stackrel{(2.10)}{>} \sum_{i=1}^{p-1} \eta_i + \sum_{i=1}^{n-2p+1} \eta_{i+p-1} \\
 &= \sum_{i=1}^{p-1} \eta_i + \sum_{i=p}^{n-p} \eta_i = \sum_{i=1}^{n-p} \eta_i.
 \end{aligned}$$

Discussion on majorizations. The Figures 1–2 at the end of the introduction show the interconnections between various types of majorization. Here through Proposition 2.9 we begin a discussion of some of these interconnections. Next we exhibit a relationship between strong majorization \preceq (recall Definition 1.3) and ∞ -majorization. As stated, it may seem to apply to summable sequences, but in fact the hypotheses, majorization and not strong majorization, negate that possibility as addressed just after the proof of Proposition 2.7.

Proposition 2.7. *If $\xi, \eta \in c_0^+$ and $\xi \prec \eta$, then*

$$\xi \not\prec \eta \implies \xi \prec_\infty \eta.$$

Proof. It suffices to show that if $\xi \prec \eta$, then the \liminf condition implies $\xi \prec_p \eta$ for every $p \in \mathbb{N}$. Indeed, suppose that

$$\liminf_{n \rightarrow \infty} \sum_{k=1}^n (\eta_k^* - \xi_k^*) = \epsilon > 0.$$

Then since $\xi \in c_0^+$ one can choose $N \in \mathbb{N}$ for which $\xi_k^* < \frac{\epsilon}{2p}$ for all $k \geq N$. Then for all n sufficiently large for which both $n \geq N$ and $\sum_{k=1}^n (\eta_k^* - \xi_k^*) > \frac{\epsilon}{2}$ one has that

$$\sum_{k=1}^{n+p} \xi_k^* < \sum_{k=1}^n \xi_k^* + p \cdot \frac{\epsilon}{2p} < \sum_{k=1}^n \eta_k^*. \quad \blacksquare$$

Note that Proposition 2.7 applies only to $\xi, \eta \notin \ell^1$, since if either ξ or η is summable, majorization implies both are summable, and in this case majorization and strong majorization are equivalent (see Definition 1.3, succeeding comment). Furthermore, the converse of Proposition 2.7 fails because there exist sequences $\xi, \eta \in c_0^* \setminus \ell^1$ for which $\xi \prec_\infty \eta$ hence also $\xi \prec \eta$, and yet $\xi \not\prec \eta$, which we exhibit in the following example.

Example 2.8. In this exposition so far we have not considered any hands-on examples of ∞ -majorization. In refuting this possible converse of Proposition 2.7 we provide one, but first we explain our natural motivation for it.

Motivation. Suppose $A \in K(H)^+$ has $\text{Tr } R_A = \infty = \text{Tr } R_A^\perp$. Then there is some basis with respect to which $A = \text{diag} \langle s_1(A), 0, s_2(A), 0, \dots \rangle$. It is a natural question to ask if there exists some $B \in E(\mathcal{U}(A))$ with $\text{Tr } R_B^\perp = 0$. The answer is yes by Theorem 2.4, but there is a more straightforward way to see this in this case. Indeed, if we let R_j denote the 2×2 rotation by $\pi/4$ acting on the subspace $\text{span}\{e_{2j-1}, e_{2j}\}$, and $U = \bigoplus R_j$, then $B := E(UAU^*) = \text{diag } \frac{1}{2}D_2s(A)$, where $D_2\varphi$ denotes the 2-ampliation of $\varphi \in \ell^\infty$, i.e., $D_2\varphi = \langle \varphi_1, \varphi_1, \varphi_2, \varphi_2, \dots \rangle$.

Example. Fix any $\eta \in c_0^*$ and choose $\xi := \frac{1}{2}D_2\eta \prec_\infty \eta$, where ∞ -majorization is easily verified (hint: $N_p = p$). Then

$$\sum_{j=1}^k (\eta_j - \xi_j) = \begin{cases} \sum_{j=n+1}^{2n} \eta_j & \text{if } k = 2n, n \in \mathbb{N}; \\ \sum_{j=n+1}^{2n-1} \eta_j + \frac{1}{2}\eta_n & \text{if } k = 2n-1, n \in \mathbb{N}. \end{cases}$$

From this it follows that

$$(2.11) \quad \left\lfloor \frac{k}{2} \right\rfloor \eta_k \leq \sum_{j=1}^k (\eta_j - \xi_j) \leq \left\lceil \frac{k}{2} \right\rceil \eta_{\lceil k/2 \rceil}.$$

From the second inequality in (2.11) it follows that if $\liminf k\eta_k = 0$, then $\xi = \frac{1}{2}D_2\eta \preceq \eta$ (e.g. for $\eta = \langle ((k+1)\log(k+1))^{-1} \rangle$, $\lim k\eta_k = 0$). However, the first inequality of (2.11) shows the inverse, that if $\liminf k\eta_k > 0$, then $\xi = \frac{1}{2}D_2\eta \not\prec \eta$ (e.g. for $\eta = \langle k^{-1} \rangle$, $\lim k\eta_{2k} = \frac{1}{2}$). The first example provides the example failing the converse of Proposition 2.7. Moreover one has, for $\eta \in c_0^*$, $\frac{1}{2}D_2\eta \preceq \eta$ if and only if $\liminf k\eta_k = 0$.

The previous example shows that for appropriate $\eta \in c_0^*$, there exist $\xi \in c_0^*$ which are counterexamples to the converse of Proposition 2.7. However, with more work one can show for *every* sequence $\eta \in c_0^+$, there is some $\xi \in c_0^*$ with $\xi \prec_\infty \eta$ and $\xi \preceq \eta$, as the next proposition shows. Though it will not be used later in this paper, we present it here for completeness.

Proposition 2.9. *For every $\eta \in c_0^+$ there exists some $\xi \in c_0^*$ with $\xi \prec_\infty \eta$ and $\xi \preceq \eta$.*

Proof. Without loss of generality, we may assume $\eta \in c_0^*$ since the definitions of ∞ -majorization and strong majorization depend only on its monotonicization η^* .

If $\liminf k\eta_k = 0$, set $\xi = \frac{1}{2}D_2\eta$. By the example preceding this proposition, we have $\xi \prec_\infty \eta$ and $\xi \preceq \eta$.

It therefore suffices to assume $\liminf k\eta_k > 0$. For this we adopt the following conventions. Let juxtaposition of finite sequences denote the concatenation of those sequences. Let $\ell(\varphi)$ denote the length of a finite sequence φ . Let $\sum \varphi$ be an abbreviation of $\sum_{j=1}^{\ell(\varphi)} \varphi_j$ and $\sum_a^b \varphi$ an abbreviation for $\sum_{j=a}^b \varphi_j$. Finally, if φ is a sequence (not necessarily finite) and $a, b \in \text{dom } \varphi$ with $a \leq b$, let $\varphi|_a^b$ denote the finite subsequence $\langle \varphi_a, \varphi_{a+1}, \dots, \varphi_b \rangle$.

If $\liminf k\eta_k > 0$ and $\eta \in c_0^*$, then $\eta > 0$. To construct ξ we proceed inductively. Let $\langle \epsilon_k \rangle$ be any positive sequence converging to zero and let $N'_0 = 1$. Then let N_1 be the smallest positive integer for which $0 < \eta_{N_1} \leq \frac{\eta_{N'_0}}{2}$. Let p_1 be the largest positive integer for which $p_1\eta_{N_1} \leq \eta_{N'_0}$, i.e., $p_1 := \left\lfloor \frac{\eta_{N'_0}}{\eta_{N_1}} \right\rfloor \geq 2$ and $(p_1 + 1)\eta_{N_1} > \eta_{N'_0} = \eta_1$. Then setting

$$\xi^{(1)} = \eta|_2^{N_1} \overbrace{\langle \eta_{N_1}, \dots, \eta_{N_1} \rangle}^{p_1},$$

this choice of p_1 guarantees that

$$(2.12) \quad 0 \leq \sum_1^{N_1} \eta - \sum \xi^{(1)} = \eta_1 - p_1\eta_{N_1} < \eta_{N_1}.$$

Denote the difference between the lengths of $\xi^{(1)}$ and $\langle \eta_j \rangle_1^{N_1}$ by

$$M_1 := \ell(\xi^{(1)}) - N_1 = p_1 - 1 \geq 1$$

and then exploiting $\eta \rightarrow 0$, choose $N'_1 > \ell(\xi^{(1)})$ for which $\sum \eta|_{N'_1-M_1}^{N'_1-1} < \epsilon_1$. Then one has

$$(2.13) \quad 0 \leq \sum_1^{N'_1-1} \eta - \sum_1^{N'_1-1} \xi^{(1)} \eta|_{N_1+1}^{N'_1-1} < \eta_{N_1} + \epsilon_1$$

because, noting that the length of $\xi^{(1)}\eta|_{N_1+1}^{N'_1-1}$ is greater by M_1 than $N'_1 - 1$, one has

$$(2.14) \quad \begin{aligned} 0 &\leq \sum_1^{N'_1-1} \eta - \sum_1^{N'_1-1} \xi^{(1)} \eta|_{N_1+1}^{N'_1-1} = \sum_1^{N'_1-1} \eta - \sum_1^{N'_1-1} \left(\xi^{(1)} \eta|_{N_1+1}^{N'_1-1} \right) + \sum \eta|_{N'_1-M_1}^{N'_1-1} \\ &= \sum_1^{N_1} \eta - \sum \xi^{(1)} + \sum \eta|_{N'_1-M_1}^{N'_1-1} \\ &< \eta_{N_1} + \epsilon_1. \end{aligned}$$

Continuing with the induction, suppose that for some $k \in \mathbb{N}$ as in the previous $k = 1$ case we are given a finite decreasing sequence $\xi^{(k)}$ and $N_k < N_k + M_k = \ell(\xi^{(k)}) < N'_k$, with $M_k \geq k$ and with the last term of $\xi^{(k)}$ being equal to η_{N_k} and $\xi^{(k)}$ satisfying

$$(2.15) \quad 0 \leq \sum_1^{N_k} \eta - \sum \xi^{(k)} < \eta_{N_k} \quad \text{and} \quad \sum \eta|_{N'_k - M_k}^{N'_k - 1} < \epsilon_k.$$

As in the previous $k = 1$ case, by an argument identical to that of (2.14), together these (2.15) inequalities imply

$$(2.16) \quad 0 \leq \sum_1^{N'_k - 1} \eta - \sum_1^{N'_k - 1} \xi^{(k)} \eta|_{N_k + 1}^{N'_k - 1} < \eta_{N_k} + \epsilon_k,$$

We then construct an extension $\xi^{(k+1)}$ of $\xi^{(k)}$ and $N_{k+1}, M_{k+1}, N'_{k+1}$ with $N'_k < N_{k+1}$ satisfying all of the properties in the preceding paragraph for $k+1$ replacing k in all instances. The procedure mimics the base case as follows. Let N_{k+1} be the smallest integer greater than N'_k for which $0 < \eta_{N_{k+1}} \leq \frac{\eta_{N'_k}}{2}$. Along with (2.15) this inequality implies that

$$\sum \xi^{(k)} + 2\eta_{N_{k+1}} \leq \sum_1^{N_k} \eta + \eta_{N'_k}.$$

Hence letting p_{k+1} be the largest positive integer for which

$$(2.17) \quad \sum \xi^{(k)} + p_{k+1}\eta_{N_{k+1}} \leq \sum_1^{N_k} \eta + \eta_{N'_k},$$

one has $p_{k+1} \geq 2$. And by the maximality of p_{k+1} one also has

$$(2.18) \quad \sum \xi^{(k)} + (p_{k+1} + 1)\eta_{N_{k+1}} > \sum_1^{N_k} \eta + \eta_{N'_k}.$$

Adding to both sides of (2.17) and (2.18) the η -terms from $N_k + 1$ to N_{k+1} excluding N'_k , one defines $\xi^{(k+1)}$ as denoted and obtains

$$(2.19) \quad \underbrace{\sum \xi^{(k)} \eta|_{N_{k+1}}^{N'_k - 1} \eta|_{N'_k + 1}^{N_{k+1}}}_{\xi^{(k+1)}} \overbrace{\langle \eta_{N_{k+1}}, \dots, \eta_{N_{k+1}} \rangle}^{p_{k+1} \text{ times}} \leq \sum_1^{N_{k+1}} \eta,$$

and

$$(2.20) \quad \sum \xi^{(k+1)} + \eta_{N_{k+1}} > \sum_1^{N_{k+1}} \eta.$$

So from (2.19)–(2.20) the difference of the (2.20) sums is nonnegative and less than $\eta_{N_{k+1}}$. This shows that $\xi^{(k+1)}$, as defined in (2.19), satisfies the first inequalities of (2.15). Note further that $\xi^{(k+1)}$ is decreasing since $\xi^{(k)}$ and η are decreasing and because the last term of $\xi^{(k)}$ is η_{N_k} .

As for M_1 , and recalling that $p_{k+1} \geq 2$, set

$$\begin{aligned} M_{k+1} &:= \ell(\xi^{(k+1)}) - N_{k+1} = (\ell(\xi^{(k)}) + N_{k+1} - (N_k + 1) + p_{k+1}) - N_{k+1} \\ &= M_k - 1 + p_{k+1} > M_k \geq k, \end{aligned}$$

and hence $M_{k+1} \geq k+1$. Next, since $\eta \in c_0^*$ we may choose some $N'_{k+1} > \ell(\xi^{(k+1)})$ satisfying the last inequality of (2.15), for $k+1$ replacing k . These facts again imply (2.16) for $k+1$ replacing k by an argument identical to (2.14).

By induction, we have constructed $\xi^{(k)}, N_k, M_k, N'_k$ with the desired properties (i.e., \P containing inequalities (2.15) and (2.16)). Furthermore, by construction each $\xi^{(k)}$ is an extension of the preceding ones. Thus, the infinite sequence ξ given by $\xi_j := \xi_j^{(k)}$ when $1 \leq j \leq \ell(\xi^{(k)})$ is well-defined. Finally it suffices to show that $\xi \prec_\infty \eta$ and $\xi \preceq \eta$.

In order to prove $\xi \prec_\infty \eta$ it suffices to observe the following two facts. Firstly, if $1 \leq m \leq N_1$ then

$$(2.21) \quad \sum_1^m \xi = \sum_1^m \eta|_2^{N_1} \langle \eta_{N_1} \rangle \leq \sum_1^m \eta.$$

Secondly, for each $k \in \mathbb{N}$, if $N_k < m \leq N_{k+1}$ then

$$(2.22) \quad \sum_1^{m+M_k} \xi \leq \sum_1^m \eta$$

because

$$\begin{aligned} \sum_1^{m+M_k} \xi &= \sum_1^{m+M_k} \xi^{(k)} \eta|_{N_k+1}^{N'_k-1} \eta|_{N'_k+1}^{N_{k+1}} \langle \eta_{N_{k+1}} \rangle \\ &= \sum_1^{\ell(\xi^{(k)})} \xi^{(k)} + \sum_{\ell(\xi^{(k)})+1}^{m+M_k} \xi^{(k)} \eta|_{N_k+1}^{N'_k-1} \eta|_{N'_k+1}^{N_{k+1}} \langle \eta_{N_{k+1}} \rangle \\ &= \sum_1^{\ell(\xi^{(k)})} \xi^{(k)} + \sum_1^{m-N_k} \eta|_{N_k+1}^{N'_k-1} \eta|_{N'_k+1}^{N_{k+1}} \langle \eta_{N_{k+1}} \rangle \\ &\leq \sum_1^{N_k} \eta + \sum_1^{m-N_k} \eta|_{N_k+1}^{N_{k+1}} = \sum_1^m \eta. \end{aligned}$$

Indeed, since $M_k \uparrow \infty$, these inequalities (2.21) and (2.22) imply $\xi \prec_k \eta$ for infinitely many $k \in \mathbb{N}$, i.e., $\xi \prec_\infty \eta$.

Finally, (2.16) implies $\xi \preceq \eta$ since

$$\liminf_{k \rightarrow \infty} \sum_{j=1}^k (\eta_j - \xi_j) \leq \liminf_{k \rightarrow \infty} \sum_{j=1}^{N'_k-1} (\eta_j - \xi_j) \leq \liminf_{k \rightarrow \infty} (\eta_{N_k} + \epsilon_k) = 0. \quad \blacksquare$$

Operator consequences. Using Theorem 2.4 and Proposition 2.7, the following corollary gives a method of ensuring membership in $E(\mathcal{U}(A))$, for $A \in K(H)^+$. The purpose of this corollary is to provide a more easily computable way to make this determination in special cases. For if one is given sequences $\xi, \eta \in c_0^+$, establishing $\xi \prec_p \eta$ or its negation seems more difficult than verifying $\xi \not\prec \eta$, which requires only $\xi \prec \eta$ and the strict positivity of the related \liminf condition.

Corollary 2.10. *Suppose $A, B \in K(H)^+$, $B \in \mathcal{D}$, $\text{Tr}(R_B^\perp) \leq \text{Tr}(R_A^\perp)$, and $s(B) \prec s(A)$ but $s(B) \not\prec s(A)$, then $B \in E(\mathcal{U}(A))$.*

Proof. By Proposition 2.7 one has $s(B) \prec_\infty s(A)$, and then using Theorem 2.4 one obtains $B \in E(\mathcal{U}(A))$. \blacksquare

3. APPROXIMATE p -MAJORIZATION (NECESSITY)

In Theorem 2.4 of the last section that if $p := \operatorname{Tr} R_A^\perp - \operatorname{Tr} R_B^\perp \geq 0$ or when undefined we set $p = 0$, then p -majorization ($s(B) \prec_p s(A)$) is a sufficient condition for membership in $E(\mathcal{U}(A))$, but it is not necessary for $0 < p < \infty$ (Example 2.6). In our quest to characterize $E(\mathcal{U}(A))$ in terms of sequence majorization, we introduce a new type of majorization called *approximate p -majorization*, which is a necessary condition for membership in $E(\mathcal{U}(A))$.

Definition 3.1. Given $\xi, \eta \in c_0^+$ and $0 \leq p < \infty$, we say that ξ is *approximately p -majorized* by η , denoted $\xi \lesssim_p \eta$, if $\xi \prec \eta$ and for every $\epsilon > 0$, there exists an $N_{p,\epsilon} \in \mathbb{N}$ such that for all $n \geq N_{p,\epsilon}$,

$$\sum_{k=1}^{n+p} \xi_k \leq \sum_{k=1}^n \eta_k + \epsilon \eta_{n+1}.$$

Furthermore, if $\xi \lesssim_p \eta$ for infinitely many $p \in \mathbb{N}$ (equivalently obviously, for all $p \in \mathbb{N}$), this we call approximate ∞ -majorization and denote it by $\xi \lesssim_\infty \eta$.

Remark 3.2. Notice from the above definition that if ξ is p -majorized by η , then ξ is trivially approximately p -majorized by η . However, there is a partial converse with a small loss in that approximate p -majorization implies $(p-1)$ -majorization. That is, if $p > 0$ and ξ is approximately p -majorized by η , then by choosing $\epsilon < 1$, from the above display ξ is $(p-1)$ -majorized by η . Combining these two facts yields that $\xi \prec_\infty \eta$ if and only if $\xi \lesssim_\infty \eta$, which is a fact we will exploit later. Furthermore, as we saw in the proof of Theorem 2.4 Case 1, if η has only finitely many nonzero terms, then any ξ which is majorized by η is ∞ -majorized by η and so also approximately ∞ -majorized by η .

Example 3.3. It is important to note that p -majorization is distinct from approximate p -majorization. That is, for each $0 < p < \infty$ we exhibit sequences $\xi, \eta \in c_0^+$ with $\xi \lesssim_p \eta$ but $\xi \not\prec_p \eta$. When $p = 1$, it suffices to consider the sequences $\xi = \langle (2^{k+1}-3)/2^{2k} \rangle$ and $\eta = \langle 2^{-k} \rangle$. Elementary calculations verify that $\xi, \eta \in c_0^*$ (that is, $\xi = \xi^*$ and $\eta = \eta^*$), and $\xi \lesssim_1 \eta$ but $\xi \not\prec_1 \eta$. To produce analogous sequences for any $p > 1$, define

$$\xi^{(p)} := \langle \overbrace{1, \dots, 1}^{p-1 \text{ times}}, \xi_1, \xi_2, \dots \rangle \quad \text{and} \quad \eta^{(p)} := \langle p-1, \eta_1, \eta_2, \dots \rangle.$$

Then $\xi^{(p)} \lesssim_p \eta^{(p)}$ but $\xi^{(p)} \not\prec_p \eta^{(p)}$, which proves that p -majorization and approximate p -majorization are distinct. However, it should be noted that these examples were not nearly as easy for us to come by as those for p -majorization. In particular the examples immediately preceding Remark 2.3 came naturally, but the single example above took some effort. For further discussion on pairs of sequences $\xi \lesssim_p \eta$ but $\xi \not\prec_p \eta$ see unifying Remark 3.7.

Our main theorem on necessity for membership in $E(\mathcal{U}(A))$ depends on approximate p -majorization:

Theorem 3.4. Suppose $A \in K(H)^+$ and $B \in E(\mathcal{U}(A))$. If

$$p = \min\{n \in \mathbb{N} \cup \{0, \infty\} \mid \operatorname{Tr}(R_A^\perp) \leq \operatorname{Tr}(R_B^\perp) + n\},$$

then $s(B) \lesssim_p s(A)$.

Proof. Suppose A, B and p are as in the hypotheses of this theorem. We may assume that A has infinite rank for otherwise the conclusion holds because of Remark 3.2 and Theorem 2.4 proof of Case 1.

We may also assume that $N := \text{Tr}(R_B^\perp) < \infty$. Otherwise, $\text{Tr}(R_B^\perp) = \infty = \text{Tr}(R_A^\perp)$ by Proposition 2.1 and therefore $p = 0$. Thus we would need to prove $s(B) \lesssim_0 s(A)$ which is equivalent to $s(B) \prec s(A)$, and this holds by Theorem 1.8.

Note that since $\text{Tr } R_B^\perp < \infty$, p satisfies $\text{Tr } R_A^\perp = \text{Tr } R_B^\perp + p = N + p$.

Case 1: $p < \infty$.

Because $\text{Tr}(R_A^\perp) = N + p$, without loss of generality via unitary equivalence for A and permutations for B , and the fact that conjugation by a permutation commutes with the expectation E , we may assume that

$$A = \text{diag}(\underbrace{0, \dots, 0}_{N+p}, s(A)) = \text{diag } \eta \quad \text{and} \quad B = \text{diag}(\underbrace{0, \dots, 0}_N, s(B)) = \text{diag } \xi.$$

Since $B \in E(\mathcal{U}(A))$ by Lemma 1.6(iv), there exists a unistochastic matrix $Q = (q_{ij})$ for which $Q\eta = \xi$. However, only double stochasticity of Q is used here. For all $m \in \mathbb{N}$ one has

$$(3.1) \quad \sum_{i=1}^m \xi_i = \sum_{i=1}^m (Q\eta)_i = \sum_{i=1}^m \sum_{j=1}^{\infty} q_{ij} \eta_j = \sum_{j=1}^{\infty} \eta_j \sum_{i=1}^m q_{ij}.$$

For a doubly-stochastic matrix $Q = (q_{ij})$, denote the last quantity in equation 3.1 as $f_m(Q, \eta)$. It is clear that for fixed η , $f_m(Q, \eta)$ depends solely on the columns of Q .

Now fix any $0 < \epsilon < 1$, and choose $N + p < N_{p,\epsilon} \in \mathbb{N}$ for which

$$(3.2) \quad N + p - \epsilon < \sum_{i=1}^{N_{p,\epsilon}} \sum_{j=1}^{N+p} q_{ij}.$$

The existence of $N_{p,\epsilon}$ follows from column-stochasticity, which yields

$$\sum_{i=1}^{\infty} \sum_{j=1}^{N+p} q_{ij} = N + p.$$

Certainly inequality (3.2) holds with $N_{p,\epsilon}$ replaced by any $m \geq N_{p,\epsilon}$, since Q is a doubly stochastic matrix and so its entries are nonnegative.

This yields an upper bound for $f_m(Q, \eta)$ for $m \geq N_{p,\epsilon}$:

$$(3.3) \quad f_m(Q, \eta) = \sum_{j=1}^{\infty} \eta_j \sum_{i=1}^m q_{ij}$$

$$(3.4) \quad = \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \sum_{j=m+1}^{\infty} \eta_j \sum_{i=1}^m q_{ij}$$

$$(3.5) \quad \leq \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \eta_{m+1} \sum_{j=m+1}^{\infty} \sum_{i=1}^m q_{ij}$$

$$(3.6) \quad = \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \eta_{m+1} \left(m - \sum_{j=1}^m \sum_{i=1}^m q_{ij} \right)$$

$$(3.7) \quad = \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \eta_{m+1} \left(\sum_{j=1}^m \left(1 - \sum_{i=1}^m q_{ij} \right) \right)$$

$$(3.8) \quad = \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \left(\sum_{j=N+p+1}^m \eta_{m+1} \left(1 - \sum_{i=1}^m q_{ij} \right) \right) \\ + \left(\sum_{j=1}^{N+p} \eta_{m+1} \left(1 - \sum_{i=1}^m q_{ij} \right) \right)$$

$$(3.9) \quad \leq \sum_{j=N+p+1}^m \eta_j \sum_{i=1}^m q_{ij} + \sum_{j=N+p+1}^m \eta_j \left(1 - \sum_{i=1}^m q_{ij} \right) \\ + \eta_{m+1} \left(N + p - \sum_{j=1}^{N+p} \sum_{i=1}^m q_{ij} \right)$$

$$(3.10) \quad < \sum_{j=N+p+1}^m \eta_j + \epsilon \eta_{m+1} \quad \text{by (3.2).}$$

Therefore, for all $m \geq N_{p,\epsilon}$ one obtains, by the definition of ξ, η ,

$$(3.11) \quad \sum_{i=1}^{m-N} s_i(B) \leq \epsilon s_{m-N-p+1}(A) + \sum_{i=1}^{m-N-p} s_i(A),$$

Since ϵ is arbitrary, this verifies Definition 3.1, that $s(B) \lesssim_p s(A)$.

Case 2: $p = \infty$.

That is, $\text{Tr } R_A^\perp = \infty$ and $N := \text{Tr } R_B^\perp < \infty$. This case is almost exactly like Case 1, except for a few minor changes. It suffices to show $s(B) \lesssim_{p'} s(A)$ for all $p' \in \mathbb{N}$. First, let $p' \in \mathbb{N}$ play the role that p played in Case 1. Now, η is $s(A)$ interspersed with infinitely many zeros (η cannot be monotonized). Furthermore, we may assume that $N + p'$ of these zeros are positioned at the beginning of the sequence η . And as in Case 1, we may assume via permutations that

$$B = \text{diag}(\underbrace{0, \dots, 0}_N, s(B)) = \text{diag } \xi.$$

Since $B \in E(\mathcal{U}(A))$, by Lemma 1.6(iv), $Q\eta = \xi$ for some unistochastic Q , and so Q is doubly stochastic. We choose $N_{p',\epsilon}$ as before, and we again want to compute $f_m(Q, \eta)$ when $m \geq N_{p',\epsilon}$, but inequalities (3.3)–(3.10) fail for this case (e.g., (3.5)). The modification trick to make them work is as follows. Consider a permutation matrix Π_m which fixes the first $N + p'$ coordinates and has the property that

$$\Pi_m \eta = \eta' = (\underbrace{0, \dots, 0}_{N+p'}, \underbrace{s_1(A), s_2(A), \dots, s_{m-N-p'+1}(A)}_{m-N-p'+1}, \eta'_{m+2}, \eta'_{m+3}, \dots),$$

where $\eta'_{m+2}, \eta'_{m+3}, \dots$ are the remainder of the $s(A)$ -terms with infinitely many zeros interspersed throughout. Then notice that Q is both doubly stochastic and satisfies inequality (3.2) if and only if $Q\Pi_m^{-1}$ does the same. Notice also that inequality (3.2) depends only on the first $N + p'$ columns of Q . Then direct computations show

$$f_m(Q, \eta) = f_m(Q\Pi_m^{-1}\Pi_m, \eta) = f_m(Q\Pi_m^{-1}, \Pi_m\eta) = f_m(Q\Pi_m^{-1}, \eta').$$

Furthermore, the computations in equations (3.3)–(3.10) hold by observation in this setting after replacing Q, η by $Q\Pi_m^{-1}, \eta'$. Therefore

$$\begin{aligned} \sum_{i=1}^{m-N} s_i(B) &= \sum_{i=1}^m \xi_i = f_m(Q, \eta) = f_m(Q\Pi_m^{-1}, \eta') \\ &\leq \sum_{j=N+p'+1}^m \eta'_j + \epsilon \eta'_{m+1} \quad \text{by (3.3)–(3.11)} \\ &= \sum_{i=1}^{m-N-p} s_i(A) + \epsilon s_{m-N-p+1}(A). \end{aligned}$$

Since ϵ is arbitrary, $s(B) \lesssim_{p'} s(A)$, and since p' is arbitrary, $s(B) \lesssim_{\infty} s(A)$. \blacksquare

One of our main results is Corollary 3.5 which, in the rather general setting where A has infinite rank and infinite dimensional kernel, we obtain a precise characterization of $E(\mathcal{U}(A))$ in terms of majorization and ∞ -majorization.

Corollary 3.5. *Suppose $A \in K(H)^+$ has infinite rank and infinite dimensional kernel ($\text{Tr } R_A = \infty = \text{Tr } R_A^{\perp}$). Then*

$$E(\mathcal{U}(A)) = E(\mathcal{U}(A))_{fk} \sqcup E(\mathcal{U}(A))_{ik},$$

the members of $E(\mathcal{U}(A))$ with finite dimensional kernel and infinite dimensional kernel, respectively, are characterized by

$$E(\mathcal{U}(A))_{fk} = \{B \in \mathcal{D} \cap K(H)^+ \mid s(B) \prec_{\infty} s(A) \quad \text{and} \quad \text{Tr } R_B^{\perp} < \infty\}$$

and

$$E(\mathcal{U}(A))_{ik} = \{B \in \mathcal{D} \cap K(H)^+ \mid s(B) \prec s(A) \quad \text{and} \quad \text{Tr } R_B^{\perp} = \infty\}.$$

Proof. If $B \in E(\mathcal{U}(A))$, then $B \in E(\mathcal{V}(A))$ and we know that $s(B) \prec s(A)$ by Theorem 1.8. But when $\text{Tr } R_B^{\perp} < \infty$, from Theorem 3.4 we know that $s(B) \lesssim_{\infty} s(A)$ which is equivalent to $s(B) \prec_{\infty} s(A)$. Thus the left-hand set in Corollary 3.5 is contained in the right-hand set.

Next suppose that $B \in \mathcal{D} \cap K(H)^+$ lies in the right-hand set. If $\text{Tr}(R_B^{\perp}) = \infty$, then Theorem 2.4 with $p = 0$ shows that $B \in E(\mathcal{U}(A))$. Similarly, if $\text{Tr}(R_B^{\perp}) < \infty$, then $s(B) \prec_{\infty} s(A)$ and again, by Theorem 2.4 for $p = \infty$, we find that $B \in E(\mathcal{U}(A))$. \blacksquare

Corollary 3.5 can be expressed as

$$E(\mathcal{U}(A)) =$$

$$\{B \in \mathcal{D} \cap K(H)^+ \mid \exists 0 \leq p \leq \infty \text{ for which } \text{Tr } R_B^{\perp} + p = \infty \text{ and } s(B) \lesssim_p s(A)\}.$$

This motivates the following conjectured characterization for $E(\mathcal{U}(A))$, which remains an open problem. And if this conjecture should prove false, is there a proper majorization characterization of $E(\mathcal{U}(A))$?

Conjecture 3.6. Let $A \in K(H)^+$. Then

$$E(\mathcal{U}(A)) =$$

$$\bigcup_{0 \leq p \leq \infty} \{B \in \mathcal{D} \cap K(H)^+ \mid s(B) \lesssim_p s(A) \text{ and } \text{Tr}(R_B^{\perp}) \leq \text{Tr}(R_A^{\perp}) \leq \text{Tr}(R_B^{\perp}) + p\}$$

The following remark provides a method for producing pairs of sequences $\xi \in c_0^+$ and $\eta \in c_0^*$ such that $\xi \lesssim_p \eta$ but $\xi \not\prec_p \eta$.

Remark 3.7. Recall that Example 2.6 provided an orthostochastic matrix Q such that for every $\eta \in c_0^*$, setting $\xi = Q\tilde{\eta}$, where $\tilde{\eta} = \langle 0, \eta \rangle$, yields $\xi \not\prec_1 \eta$. Using Lemma 1.6 one has $\text{diag } \xi \in E(\mathcal{U}(\text{diag } \eta))$. By Theorem 3.4,

$$\xi^* = s(\text{diag } \xi) \lesssim_1 s(\text{diag } \tilde{\eta}) = \tilde{\eta}^* = \eta,$$

hence $\xi \lesssim_1 \eta$. In general, given two finite sequences, say φ, ζ , of lengths n and $n+p$ (where n is arbitrary) and having the same sum (i.e., $\sum_{j=1}^n \varphi_j = \sum_{j=1}^{n+p} \zeta_j$), one can prepend φ to η and ζ to ξ , clearly obtaining new sequences ξ', η' with $\xi' \lesssim_p \eta'$ but $\xi' \not\prec_p \eta'$.

Although we did not mention it earlier in Example 2.6, there is a significant amount of freedom in choosing Q . In particular, examination of [KW10, Example 6.11] ensures that each sum-1 strictly positive column vector followed by a Gram-Schmidt process produces a distinct orthostochastic matrix Q that can be used in Example 2.6. Perhaps these orthostochastic Q can be exploited or modified to prove Conjecture 3.6.

4. $E(\mathcal{U}(A))$ CONVEXITY

Historically, convexity played a central role and is ubiquitous in majorization theory. For example, a theorem of Hardy, Littlewood and Pólya [HLP88] proves

$$\{\xi \in \mathbb{R}^n \mid \xi \prec \eta\} = \text{co}\{\tilde{\eta} \in \mathbb{R}^n \mid \tilde{\eta}_k = \eta_{\pi(k)}, \pi \in \Pi_n\},$$

where Π_n is the set of $n \times n$ permutation matrices. For operators, in [Hor54], using [Sch23], Horn proved that $E(\mathcal{U}(X))$ is convex whenever $X = X^* \in M_n(\mathbb{C})$ by establishing the characterization

$$E(\mathcal{U}(X)) = \{\text{diag } d \mid d \in \mathbb{R}^n, d \prec \lambda\},$$

where λ is the eigenvalue sequence of X (Theorem 1.1). However, the verification that $E(\mathcal{U}(X))$ is convex is immediate from its majorization characterization even without the theorem of Hardy, Littlewood and Pólya mentioned above. Likewise it is straightforward to verify that if $\eta \in c_0^+$, then

$$\{\xi \in c_0^+ \mid \xi \prec \eta\}$$

is convex. In particular, this leads to the results of Kaftal and Weiss on the convexity of the expectation of the partial isometry orbit of a positive compact operator.

Theorem 4.1 ([KW10, Corollary 6.7]). *Let $A \in K(H)^+$. Then*

- $E(\mathcal{V}(A))$ is convex.
- If $R_A = I$ or A has finite rank, then $E(\mathcal{U}(A))$ is convex.

Since we have a characterization of $E(\mathcal{U}(A))$ when $A \in K(H)^+$ has both infinite rank and infinite dimensional kernel, it seems natural to ask if $E(\mathcal{U}(A))$ is convex in this case. The answer is positive, however the verification is much less obvious to us (see below Corollary 4.3). But first, a lemma.

Lemma 4.2. *Suppose that $\xi, \zeta, \eta \in c_0^+$, $0 < \lambda < 1$, $0 \leq p, q \leq \infty$ such that $\xi \lesssim_p \eta$, $\zeta \lesssim_q \eta$. If $r = \min\{p + |\xi^{-1}(0) \setminus \zeta^{-1}(0)|, q + |\zeta^{-1}(0) \setminus \xi^{-1}(0)|\}$, then $\lambda\xi + (1-\lambda)\zeta \lesssim_r \eta$.*

Proof. Set $\varphi := \lambda\xi + (1 - \lambda)\zeta$. There are two cases: either η has finite support, or not. If the former, then since $\xi \prec \eta$ and $\zeta \prec \eta$, one easily has $\varphi \prec \eta$ by the comment immediately preceding Theorem 4.1. Then since η has finite support, we can improve this to $\varphi \prec_\infty \eta$ (see Remark 3.2 or proof of Theorem 2.4 Case 1), which is equivalent to $\varphi \lesssim_\infty \eta$. Thus $\varphi \lesssim_r \eta$.

The second case: η has infinite support. For now, suppose both p, q are finite. Let $\pi : \mathbb{N} \rightarrow \mathbb{N} \setminus \varphi^{-1}(0)$ be a bijection monotonicizing φ : $\varphi_k^* = \varphi_{\pi(k)}$. Since $\varphi^{-1}(0) = \xi^{-1}(0) \cap \zeta^{-1}(0)$, one has

$$(4.1) \quad \mathbb{N} \setminus \varphi^{-1}(0) = (\mathbb{N} \setminus (\xi^{-1}(0) \cup \zeta^{-1}(0))) \sqcup (\xi^{-1}(0) \setminus \zeta^{-1}(0)) \sqcup (\zeta^{-1}(0) \setminus \xi^{-1}(0)),$$

which immediately yields the disjoint union

$$\begin{aligned} \pi([1, m]) &= [\pi([1, m]) \cap (\mathbb{N} \setminus (\xi^{-1}(0) \cup \zeta^{-1}(0)))] \text{ (cardinality } k_m) \\ &\quad \sqcup [\pi([1, m]) \cap (\xi^{-1}(0) \setminus \zeta^{-1}(0))] \text{ (cardinality } s_m) \\ &\quad \sqcup [\pi([1, m]) \cap (\zeta^{-1}(0) \setminus \xi^{-1}(0))] \text{ (cardinality } t_m) \end{aligned}$$

for each $m \in \mathbb{N}$. Therefore $m = k_m + s_m + t_m$, and each term increases to the cardinality of its corresponding set from equation (4.1). From this cardinality equation, it is clear that $m - s_m = k_m + t_m$ and $m - t_m = k_m + s_m$ are both increasing sequences. However, we may further conclude that they increase without bound. Indeed,

$$\begin{aligned} m - s_m &= |\pi([1, m]) \cap (\mathbb{N} \setminus (\xi^{-1}(0) \cup \zeta^{-1}(0)))| \\ &\quad + |\pi([1, m]) \cap (\xi^{-1}(0) \setminus \zeta^{-1}(0))| \\ &= |\pi([1, m]) \cap (\mathbb{N} \setminus \zeta^{-1}(0))|, \end{aligned}$$

which shows $m - s_m \uparrow |\mathbb{N} \setminus \zeta^{-1}(0)| = |\text{supp } \zeta|$. Likewise $m - t_m \uparrow |\text{supp } \xi|$. And ξ, ζ are both infinitely supported since they are majorized by η and η is infinitely supported (as shown simply in the proof of Theorem 2.4 Case 2, first ¶). Thus we have verified $m - s_m, m - t_m \uparrow \infty$. Therefore, given $0 < \epsilon < 1$, once m is sufficiently large so that $m - s_m \geq N_{p, \epsilon}$ and $m - t_m \geq N_{q, \epsilon}$, one has

$$\begin{aligned} \sum_{k=1}^m \varphi_k^* &= \lambda \sum_{k=1}^m \xi_{\pi(k)} + (1 - \lambda) \sum_{k=1}^m \zeta_{\pi(k)} \\ &\leq \lambda \sum_{k=1}^{m-s_m} \xi_k^* + (1 - \lambda) \sum_{k=1}^{m-t_m} \zeta_k^* \\ &\leq \lambda \left(\sum_{k=1}^{m-s_m-p} \eta_k^* + \epsilon \eta_{m-s_m-p+1} \right) \\ &\quad + (1 - \lambda) \left(\sum_{k=1}^{m-t_m-q} \eta_k^* + \epsilon \eta_{m-t_m-q+1}^* \right) \\ &\leq \sum_{k=1}^{m-r_m} \eta_k^* + \epsilon \eta_{m-r_m+1}^*, \end{aligned}$$

where $r_m = \min\{s_m + p, t_m + q\}$. The above computation proves that $\varphi \lesssim_{r_m} \eta$. But since $r_m \uparrow r$, either $r_m \uparrow \infty$, or eventually $r_m = r$. In either case $\varphi \lesssim_r \eta$.

Finally, to remove the restriction that p, q be finite, observe that the above proof actually showed that for any $p' \leq p, q' \leq q$ with $p', q' < \infty$, one has $\varphi \lesssim_{r'_m} \eta$, where $r'_m = \min\{s_m + p', t_m + q'\}$. The proof now splits into several subcases.

If $p < \infty, q = \infty$ and $|\xi^{-1}(0) \setminus \zeta^{-1}(0)| < \infty$, then choosing $p' = p$ and $q' = p + |\xi^{-1}(0) \setminus \zeta^{-1}(0)|$, one has $r'_m = r = p + |\xi^{-1}(0) \setminus \zeta^{-1}(0)|$ for m sufficiently large so that $s_m = |\xi^{-1}(0) \setminus \zeta^{-1}(0)|$.

If $p < \infty, q = \infty$ and $|\xi^{-1}(0) \setminus \zeta^{-1}(0)| = \infty$, then for $p' = p$ and any $q' < \infty$, eventually r'_m reaches q' since $s_m \uparrow \infty$. Therefore $\varphi \lesssim_{q'} \eta$, which means $\varphi \lesssim_\infty \eta$ since q' was arbitrary.

The cases where $p = \infty$ and $q < \infty$ hold by symmetric arguments.

If $p = \infty = q$, then for $p' = q' = k \in \mathbb{N}$, one has $r'_m \geq k$ and so $\varphi \lesssim_k \eta$, hence $\varphi \lesssim_\infty \eta$. ■

Examination of the proof of Lemma 4.2 actually shows that we may replace approximate p -majorization with p -majorization everywhere in the statement of the lemma and the result remains valid. Indeed, the only difference in the proof is that the terms involving ϵ disappear when p -majorization is used.

An operator $E(\mathcal{U}(A))$ consequence of this is:

Corollary 4.3. *If $A \in K(H)^+$ and $\text{Tr}(R_A) = \infty = \text{Tr}(R_A^\perp)$, then $E(\mathcal{U}(A))$ is convex.*

Proof. Take $B, C \in E(\mathcal{U}(A))$ and $0 < \lambda < 1$. Let $D = \lambda B + (1 - \lambda)C$, and let $d = \lambda b + (1 - \lambda)c$ be their corresponding diagonal sequences in c_0^+ . If $\text{Tr}(R_D^\perp) = \infty$, that $D \in E(\mathcal{U}(A))$ follows immediately from Corollary 3.5 and from the fact that convex combinations of elements majorized by η are themselves majorized by η . However, if $\text{Tr}(R_D^\perp) < \infty$, we use the dichotomy of Corollary 3.5 for B, C . Observing that $\text{Tr}(R_B^\perp) = |b^{-1}(0)|$ and similarly for C, D , simply notice that either $|b^{-1}(0) \setminus d^{-1}(0)| = \infty$ or $|b^{-1}(0) \setminus d^{-1}(0)| < \infty$ (so $\text{Tr} R_B^\perp < \infty$), in which case $b \prec_\infty a$ by Corollary 3.5. Likewise, $|c^{-1}(0) \setminus d^{-1}(0)| = \infty$ or $c \prec_\infty a$. Thus by Lemma 4.2 one has $d \prec_\infty a$. And so by Corollary 3.5, $D \in E(\mathcal{U}(A))$. ■

REFERENCES

- [AK06] W. Arveson and R. V. Kadison, *Diagonals of self-adjoint operators*, Operator Theory, Operator Algebras, and Applications **414** (2006), 247–263.
- [AM07] M. Argerami and P. Massey, *A Schur-Horn theorem in II_1 factors*, Indiana Univ. Math. J. **56** (2007), no. 5, 2051–2059. MR 2359722 (2008m:46120)
- [AM08] ———, *A contractive version of a Schur-Horn theorem in II_1 factors*, J. Math. Anal. Appl. **337** (2008), no. 1, 231–238. MR 2356067 (2008m:46121)
- [AM13] ———, *Schur-Horn theorems in II_∞ -factors*, Pacific J. Math. **261** (2013), no. 2, 283–310. MR 3037568
- [BJ12] M. Bownik and J. Jasper, *Diagonals of self-adjoint operators with finite spectrum*, Preprint, arXiv:1302.5106 [math.FA], June 2012.
- [BJ13] ———, *The Schur-Horn Theorem for operators with finite spectrum*, Preprint, arXiv:1302.4757 [math.FA], February 2013.
- [GM64] I. C. Gohberg and A. Markus, *Some relations between eigenvalues and matrix elements of linear operators*, Mat. Sb. (N.S.) **64** (106) (1964), 481–496. MR 0170218 (30 #457)
- [HLP88] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities. 2nd ed., 1st. paperback ed., 2nd ed., 1st. paperback ed ed., Cambridge (UK) etc.: Cambridge University Press, 1988* (English).
- [Hor54] A. Horn, *Doubly stochastic matrices and the diagonal of a rotation matrix*, American Journal of Mathematics **76** (1954), 620–630.

- [KS08] N. J. Kalton and F. Sukochev, *Symmetric norms and spaces of operators.*, Journal für die Reine und Angewandte Mathematik **621** (2008), 81–121 (English).
- [KW10] V. Kaftal and G. Weiss, *An infinite dimensional Schur-Horn Theorem and majorization theory*, Journal of Functional Analysis **259** (2010), no. 12, 3115–3162.
- [Mar64] A. Markus, *Eigenvalues and singular values of the sum and product of linear operators*, Uspehi Mat. Nauk **19** (1964), no. 4 (118), 93–123. MR 0169063 (29 #6318)
- [Neu99] A. Neumann, *An infinite dimensional version of the Schur-Horn convexity theorem*, Journal of Functional Analysis **161** (1999), no. 2, 418–451.
- [Rav12] M. Ravichandran, *The Schur-Horn Theorem in von Neumann algebras*, Preprint, arXiv:1209.0909 [math.OA], September 2012.
- [Sch23] I. Schur, *Über eine klasse von mittelbildungen mit anwendungen auf der determinantentheorie*, Sitzungsber. Berliner Mat. Ges. **22** (1923), 9–29.

(Jireh Loreaux) UNIVERSITY OF CINCINNATI, DEPARTMENT OF MATHEMATICS, CINCINNATI, OH, 45221-0025, USA

E-mail address: loreaujy@mail.uc.edu, loreaujy@gmail.com

(Gary Weiss) UNIVERSITY OF CINCINNATI, DEPARTMENT OF MATHEMATICS, CINCINNATI, OH, 45221-0025, USA

E-mail address: gary.weiss@uc.edu, weissg@ucmail.uc.edu, gary.weiss@math.uc.edu